Demography, Vol. 25, No. 3, August 1988

# Multinomial and Conditional Logit Discrete-Choice Models in Demography

**Saul D. Hoffman**
Department of Economics, University of Delaware, Newark, Delaware 19716
**Greg J. Duncan**
Institute for Social Research, University of Michigan, Ann Arbor, Michigan 48106

Although discrete-choice statistical techniques have been used with increasing regularity in demographic analyses, McFadden's conditional logit model is less well known and seldom used. Conditional logit models are appropriate when the choice among alternatives is modeled as a function of the characteristics of the alternatives, rather than (or in addition to) the characteristics of the individual making the choice. We argue that this feature of conditional logit makes it more appropriate for estimating behavioral models. In this article, the conditional logit model is presented and compared with the more familiar multinomial logit model. The difference between the two techniques is illustrated with an analysis of the choice of marital and welfare status by divorced or separated women.

Statistical techniques for the analysis of discrete choices have been used with increasing regularity in demographic analyses.[1] The best known are the binomial logit and probit techniques, both of which are suitable for binary choice problems. For problems involving the choice among three or more categories, the multinomial logit technique is most often employed; the corresponding probit model is used relatively little because of its computational difficulty. Virtually unused thus far is a closely related technique called conditional logit, a model that is well suited for behavioral modeling of polychotomous choice situations. Developed by McFadden (1973), the conditional logit model is widely used in transportation demand studies (see Ben-Akiva and Lerman, 1985) but is seldom used in demographic research.[2]

Conditional logit is not simply a different and arguably preferable technique for estimating the kind of models for which multinomial logit is currently used. Rather, it is appropriate for a different class of models in which a choice among alternatives is treated as a function of the characteristics of the alternatives, rather than (or in addition to) the characteristics of the individual making the choice.

We believe that many problems of interest to demographers and other social scientists can be modeled by using a "characteristics of the alternative" approach. Thus they are appropriately estimated with conditional logit. Furthermore, we suggest that it is often difficult to attach a behavioral interpretation to the results of models that focus exclusively on the "characteristics of the chooser"—that is, those estimated by conventional multinomial logit.

The next section of this article describes the basic statistical properties of the conditional logit (CLGT) model and compares it with the better known multinomial logit (MNLGT) model.[3] It also considers the form of the underlying models of individual behavior that lead to MNLGT and CLGT estimation. The third section presents a brief discussion of some practical statistical and estimation issues relating to the CLGT model. The final section uses data from the Panel Study of Income Dynamics to illustrate the difference between the two techniques in applied work. We examine divorced or separated women's choice among a set of marital and welfare status alternatives by first using a pure MNLGT model, then a pure CLGT model, and then a mixed version that incorporates features of both.

## Statistical and Modeling Issues

Both multinomial logit and conditional logit are used to analyze the choice of an individual among a set of $J$ alternatives. The central distinction between the two can be put very simply: MNLGT focuses on the *individual* as the unit of analysis and uses the individual's characteristics as explanatory variables; in contrast, CLGT focuses on the set of *alternatives* for each individual and the explanatory variables are characteristics of those alternatives.[4]

Let $X_i$ stand for the characteristics of individual $i$ and $Z_{ij}$ for the characteristics of the $j$th alternative for individual $i$, with the corresponding parameter vectors denoted by $\beta$ and $\alpha$, respectively. Let $J$ be the number of unordered alternatives (for the moment, assumed constant for all individuals) and $P_{ij}$ the probability that individual $i$ chooses alternative $j$. The choice probabilities in the MNLGT and CLGT models are

MNLGT: $$P_{ij} = \exp(X_i\beta_j) \Big/ \sum_{k=1}^{J} \exp(X_i\beta_k), \quad (1)$$

CLGT: $$P_{ij} = \exp(Z_{ij}\alpha) \Big/ \sum_{k=1}^{J} \exp(Z_{ik}\alpha). \quad (2)$$

In a mixed model that includes both characteristics of the alternatives and the individual, the corresponding probability can be written as

Mixed: $$P_{ij} = \sum_{k=1}^{J} \exp(X_i\beta_j + Z_{ij}\beta)/\exp(X_i\beta_k + Z_{ik}\alpha). \quad (3)$$

We discuss the mixed logit model (3) further in the next section and estimate such a model in the last section of this article.[5]

Note the symmetry in equations (1) and (2). In the MNLGT model, the explanatory variables (X), being characteristics of the individual, are themselves constant across the alternatives. Consequently, the only way they can affect choice probabilities is by having a different impact on the various alternatives. Thus in practice, MNLGT estimates a set of $J - 1$ coefficients ($\beta_j$) for each explanatory variable. The estimated coefficients show the effect of the X variables on the probability of choosing each alternative relative to one alternative that serves as a common benchmark. There are only $J - 1$ coefficients, because the scaling of the coefficients is arbitrary. Thus it is necessary to normalize on one set of coefficients, typically by setting it equal to zero. For this alternative, the corresponding probability is $1/\Sigma \exp(X_{ij}\beta_j)$, since $\beta = 0$ and $\exp(0) = 1$.

In contrast, in the CLGT model, the explanatory variables (Z) assume different values in each alternative (note the presence of a $j$ subscript on Z but not X), but the impact of a unit of Z is usually, although not necessarily, assumed to be constant across alternatives. In that case, only a single coefficient is estimated for each Z variable, so the impact of a variable on the choice probabilities derives from the difference in its value across alternatives. Consequently, in the standard CLGT formulation, a Z (or X) variable with no variation across alternatives has no impact on choice probabilities. When such variables are deemed to be important, the mixed model is required.

The basic difference between the MNLGT and CLGT formulations is clearer when equations (1) and (2) are rewritten by dividing through by the numerator:

MNLGT: $$P_{ij} = 1 \Big/ \sum_{k=1}^{J} \exp[X_i(\beta_k - \beta_j)], \tag{4}$$

CLGT: $$P_{ij} = 1 \Big/ \sum_{k=1}^{J} \exp[(Z_{ik} - Z_{ij})\alpha]. \tag{5}$$

Here, the probability in equation (4) depends on the difference in the coefficients across alternatives, whereas in equation (5), the probability depends on the differences in the value of the characteristics across alternatives.

The difference between the MNLGT and CLGT models is not merely one of statistical form. The choice probabilities in equations (1) and (2) reflect the underlying models of individual behavior that necessarily reflect hypotheses about the basis on which individuals make choices among alternatives. Often this is not made explicit, and researchers move to their empirical estimation without first specifying the underlying behavioral model. In fact, however, it is a crucial step for the interpretation of the empirical results.

Let $V_{ij}$ stand for the value (utility) of alternative $j$ to individual $i$, and assume, as a behavioral rule, that an individual chooses his or her most highly valued alternative. Suppose that $V_{ij}$ depends on the attributes of the alternatives ($Z_j$) through some unspecified functional form ($f_1$). Then the choice problem can be represented by a pair of equations as follows:

$$V_{ij} = f_1(Z_{ij}), \tag{6}$$

$$P_{ij} = \Pr(V_{ij} > V_{ik}) \quad \text{all } k \text{ not equal to } j. \tag{7}$$

With the addition of an appropriately defined error term,[6] equation (6) leads to the CLGT model rather than the MNLGT model, since the characteristics of the alternatives are the determinants of choice. The estimated parameters of $f_1$ provide information not only about the choice probabilities through equation (2) but also about the value function in equation (6).

The specific form of equation (6) will, of course, vary with the nature of the problem and the discipline. Economists, for example, virtually always regard utility as a function primarily of an individual's level of consumption (defined broadly) or, equivalently, the exogenous income and the set of prices he or she faces. Viewed in this way, equation (6) is a statement about the functional relationship between the characteristics of the alternatives (the $Z_{ij}$'s) and the utility of each alternative to the individual (the $V_{ij}$'s)—in short, a utility function. Equation (7) represents the well-known principle of utility maximization applied to a discrete choice problem. We estimate a version of equation (6) in the last section of this article.

Noneconomic models based on equation (6) could also be formulated, although we know of no attempts to do so. For example, a choice model of becoming married versus remaining single might view the value of these two alternatives as a function of the economic security, companionship, independence, and other attributes that each provides, with the perceived extent of these attributes in each alternative obtained through survey questions. To avoid problems with respondents' rationalizing past decisions, a useful research design might be a two-wave panel in which the attitudinal information is ascertained in the first wave and the behavior (e.g., transition to marriage) is measured in the second. Thus reports by unmarried respondents about the attributes of the various marital states in a first wave could be used to predict the probability of marriage in a follow-up interview.

Another example would be a child-care choice model in which the value of a given child-care model (e.g., day care, sitter in own home) is taken to be a function of characteristics such as its likely effect on child development, its cost, and its reliability.

Note that in these two examples the perceived or objective *characteristics* of each alternative rather than its subjective *importance* or *satisfaction* are used to explain an individual's choice. The statistical model would then provide information (i.e., the estimated coefficients) about the relative value that individuals place on the various characteristics, inferred from their actual behavior.

What behavioral model leads to MNLGT estimation? In general, MNLGT will be appropriate when equation (6) is replaced with

$$V_{ij} = f_2(X_i), \tag{8}$$

where $f_2$ is some unspecified function relating $X_i$ to $V_{ij}$. In equation (8), the value of an alternative is regarded as a function of the characteristics of the individual. Assuming that equation (7) still holds, equation (8) then leads to MNLGT estimation.

Multinomial logit can also be shown to represent a nonbehavioral, reduced-form version of equations (6) and (7). If equation (6) holds but

$$Z_{ij} = g(X_i), \tag{9}$$

then

$$V_{ij} = h(X_i). \tag{10}$$

Equation (10) relates the value of the $j$th alternative to the characteristics of the $i$th individual, but without including the characteristics of the $j$th alternative.

In a sense, the choice between MNLGT and CLGT is the choice between a model represented by either equation (8) or equation (10) and one represented by equation (6). Although there is, of course, no general rule about which is preferable, we think that a model based on equation (6) and utilizing equation (7) has much to recommend it. The general notion that structural models are preferable to reduced-form models is one argument in favor of equation (6) and CLGT estimation instead of equation (10) and MNLGT estimation. Even though models such as equation (8) may provide direct and useful information about which individuals make which choices, they are often not well suited to testing hypotheses about why those choices are made. Indeed, the interpretation of models such as equation (8) often make reference to the (untested) characteristics of alternatives available to particular individuals.

Models such as equation (6) are especially well suited for the analysis of situations in which government policy affects the attractiveness of an alternative by changing some relevant characteristic. Examples include the Aid to Females With Dependent Children (AFDC) program, which provides income to female-headed families with children; scholarship aid for higher education, which may make college attendance more attractive; and subsidies for day care, which may increase the labor force activity of women. To assess the effect of government policies like these on individual choices, it is necessary, when possible, to include the policy parameters directly in the choice problem. Since these parameters, though, are typically a characteristic of the alternative in question, a conditional logit model such as equation (6) is the appropriate model.

## Statistical Properties and Estimation Issues

In this section, we present a brief survey of some of the practical statistical issues involved in the estimation of the MNLGT and CLGT models. Among other things, we call attention here to one of the potentially undesirable restrictions imposed by the logistic form used in either model. We also discuss some estimation issues.

*Likelihood Function.* Despite the differences discussed earlier, the CLGT and MNLGT models share a common likelihood function:

$$\log L = \sum_i \sum_j y_{ij} P_{ij}, \tag{11}$$

where $y_{ij} = 1$ if individual $i$ chooses alternative $j$ and equals 0 otherwise. The difference between the models is in the formulation of the choice probabilities, as in equations (1) and (2), and in the underlying behavioral models that they represent. As a practical matter, these differences often lead to differences in the way the data are prepared for estimation and in the software programs used to estimate the two models.

*Statistical Specification.* Both the CLGT and MNLGT models are based on the assumption that the error terms in equations (6), (8), and (10) follow an extreme value distribution and are independent across alternatives. The assumption of independence is critical; any other assumption leads to substantial computational difficulties involving the computation of multivariate integrals. The "cost" of the independence assumption is the so-called "independence of irrelevant alternatives" (IIA) problem. As derived from equation (2), the *ratio* of the choice probabilities for any two of the $j$ alternatives depends only on the characteristics of those two alternatives. If, for example, there is a change in the characteristics of any other alternative in the choice set, this property requires that the two probabilities must adjust precisely in order to preserve their initial ratio.[7] This is equivalent to assuming that the percentage change in each probability is equal, a response pattern that may be an unwarranted and inappropriate restriction. For example, the possibility that one choice probability might be more greatly affected by such a change is thereby excluded.

As a practical matter, the independence assumption is most likely to be problematic when the alternatives are similar to one another, so that unobserved factors affecting one alternative may well affect another alternative. The IIA assumption can be tested (Hausman and McFadden, 1984). If it is not supported, there are two general alternatives. One is the conditional probit model, which allows for multivariate normal correlated error terms. The other is the nested logit model (Hausman and McFadden, 1984; McFadden, 1981) in which the choice process is viewed as a set of nested choices. This approach retains the computational advantages of the logit form but selectively relaxes the independence assumption and thereby allows a variety of response patterns to a change in the characteristics of one alternative.

*Statistical Software.* Most general-purpose statistical software packages contain a bivariate logit procedure and an MNLGT procedure. Software to estimate the CLGT model is less common. Our CLGT estimation uses the Discrete Choice procedure available in LIMDEP; the MNLGT model is estimated with LIMDEP's Logit procedure. The Mlogit procedure in SAS can be used to estimate both the MNLGT and CLGT models.

*Estimation Details.*[8] The estimation of a CLGT model is somewhat unorthodox, because the unit of analysis is, in some sense, not the individual but, rather, the set of alternatives available to each individual.

Consider $N$ individuals, each of whom has $J$ alternatives. To estimate a CLGT model, an individual's record is transformed into $J$ distinct records, each one representing an alternative for that individual. The alternatives are represented in the same sequence for each individual; the first record represents alternative 1, the second alternative 2, and so on to record and alternative $J$. The explanatory variables are similarly constructed to reflect the value of each variable for each individual in each alternative. An individual's choice among the alternatives is indicated by a 1 for the appropriate record; the other alternatives are coded 0.

Table 1 illustrates the typical data structure for CLGT estimation. In this example, there are three alternatives for each of four individuals, who choose alternatives 3, 2, 1, and 2, respectively. There is one X variable and one Z variable; the inclusion of individual characteristics means that the model is really a mixed model rather than a pure CLGT model. $Z_{11}$ is the value for individual 1 of some characteristic in alternative 1, $Z_{12}$ is the value of that characteristic for that individual in alternative 2, $Z_{21}$ is the value of that characteristic for individual 2 in alternative 1, and so on. Estimation of this model would yield a single coefficient for Z.

The final two columns show how an attribute that is invariant across alternatives can be introduced to create a mixed logit model. Let $D_2$ be a dummy variable equal to 1 for alternative 2 and 0 for the other alternatives, and let $D_3$ be defined similarly for alternative 3. The variables in the final two columns are $D_2X$ and $D_3X$; just as in MNLGT estimation, they give the effect of variable X relative to an omitted category, here alternative 1. Estimation of this model would yield three coefficients—one each for Z, $XD_2$, and $XD_3$. If desired, constant terms for two alternatives could be constructed by using $D_2$ and $D_3$.

There is an additional and particularly useful feature of CLGT models. In some choice situations, not every alternative is available to every individual. For example, women without children are categorically ineligible to receive AFDC, and only women living in states offering the AFDC–UP program can choose to be both married and receiving welfare; only widows with living children can choose to live with their children; only individuals owning cars or living near bus routes can drive or take the bus to work, respectively. Taking proper account of differences in the size and composition of the choice set available to specific individuals is troublesome under most circumstances and often leads to clumsy, ad hoc solutions. The sample may be partitioned so that the analysis is no longer general, or 0 values might be assigned for the independent variables in cases like that concerning AFDC benefits of women ineligible to receive AFDC.[9] A more natural solution, however, is simply to eliminate an irrelevant alternative from the choice set for an individual.[10] With the choice

Table 1. Typical Data Structure for CLGT Estimation

| Individual | Alternative | Dependent variable | $Z$ | $XD_2$ | $XD_3$ |
|---|---|---|---|---|---|
| 1 | 1 | 0 | $Z_{11}$ | 0 | 0 |
| | 2 | 0 | $Z_{12}$ | $X_1$ | 0 |
| | 3 | 1 | $Z_{13}$ | 0 | $X_1$ |
| 2 | 1 | 0 | $Z_{21}$ | 0 | 0 |
| | 2 | 1 | $Z_{22}$ | $X_2$ | 0 |
| | 3 | 0 | $Z_{23}$ | 0 | $X_2$ |
| 3 | 1 | 1 | $Z_{31}$ | 0 | 0 |
| | 2 | 0 | $Z_{32}$ | $X_3$ | 0 |
| | 3 | 0 | $Z_{33}$ | 0 | $X_3$ |
| 4 | 1 | 0 | $Z_{41}$ | 0 | 0 |
| | 2 | 1 | $Z_{42}$ | $X_4$ | 0 |
| | 3 | 0 | $Z_{43}$ | 0 | $X_4$ |

set as the unit of observation, tailoring the choice set to individual circumstances is a straightforward matter.

## An Empirical Example: Remarriage and Welfare Choices of Divorced and Separated Women

In this section, we examine the remarriage and welfare choices of divorced or separated women. We present estimates of three models—a standard MNLGT model with individual characteristics as explanatory variables, a pure CLGT model with characteristics of the alternatives as explanatory variables, and a mixed model that includes characteristics of both the individual and the alternatives.

Our analysis is based on data from the Panel Study of Income Dynamics (PSID) on white women under the age of 45 who became divorced or separated between 1969 and 1982. Each woman is observed from the date of her divorce or separation until remarriage, the end of the panel observation period, or the sixth post-divorce/separation year, whichever comes first. Our data are in person-year event-history format, defined over a spell of being "unmarried." Formally, we are estimating a discrete-time hazard model of time until remarriage, using an MNLGT or CLGT model as the estimation procedure. Time-varying independent variables are measured as of the person-year used in the analysis. See Allison (1982, 1984) for a general discussion of discrete-time hazard models.[11]

In each year, a woman is observed in one of three alternative states: she can remarry, she can remain single and receive welfare, or she can remain single without receiving welfare. We use functional rather than legal definitions of marriage, divorce, and remarriage. Unmarried couples are treated as married by the PSID if they reside together for two consecutive interviews; given this definition, we can analyze the "remarriage" choice of separated women. Welfare receipt is defined as receiving a dollar or more of income from AFDC or the "other welfare" category used in the PSID.[12]

We are interested in analyzing the determinants of the trichotomous choice of remarriage, welfare receipt, and remaining single without welfare receipt. The motivation includes understanding the potential role of AFDC income in discouraging remarriage as well as the more general determinants of remarriage decisions.

One explanation, cast in terms of individual characteristics, might focus on such things as a woman's age and education, the number of children she has, and whether she resides in an urban area. We estimate this model with the MNLGT model.

A different explanation might consider, instead, the exogenous income (the income available to a woman at zero hours of work) and her net (after-tax) wage rate in each of the three alternatives. This corresponds to a model like equation (6) in which the value of an alternative is a function of its characteristics, here exogenous income and prices. Technically, we are using the concept of indirect utility functions in which the maximum utility (satisfaction) available to an individual in an alternative depends on its exogenous income and the set of prices (in our model, the wage rate) it provides.

Consider, first, the exogenous income available to a woman in each alternative. While some income, such as child support, and income from dividends and interest are unaffected by her choice of alternative, other components vary systematically by alternatives. For instance, if she were to accept welfare, she would receive the legally mandated benefits paid in her state of residence, given her family size and other income. If she were to marry, she would be ineligible for welfare and benefits,[13] but she would have access to some portion of her new husband's income. If she remained single without accepting welfare, she would receive alimony and/or child support income, if any, plus any income from dividends and interest.

Her after-tax wage rate would also differ across alternatives, even though the market (pretax) wage rate for a particular woman is likely to be constant across alternatives. After-tax wages are particularly low in the welfare alternative by virtue of the high benefit reduction rate applied to earned income in welfare and the existence of an earnings ceiling to maintain welfare eligibility.

This model, in which choice among alternatives is a function of the exogenous income and wage characteristics of each alternative, is estimated with the CLGT model. We also estimate a mixed model that includes the individual variables from the MNLGT model and the alternative-specific variables from the CLGT model. In both the pure CLGT and the mixed models, we allow the number of alternatives to vary across individuals. Women without dependent children and women with substantial nonlabor income are ineligible for welfare, and thus that alternative is not available to them.

The characteristics of the sample, including sample size and mean value of all of the independent variables, are presented in Table 2. The MNLGT results appear in Table 3, columns 1 and 2. The coefficients express effects relative to the omitted category, single/welfare. We find that more educated women are more likely to be either married or single/no welfare than to be receiving welfare, whereas residence in an urban area and having more children both decrease those probabilities. Older women are more likely to be single and not receive welfare, but they are no more likely to be married. Despite the statistically significant coefficients, it is not easy to explain why these variables have the impacts that they do: Do

Table 2. Characteristics of PSID Sample of White Women Undergoing Divorce or Separation, 1969–1982

| Characteristic | No. | Mean |
|---|---|---|
| Sample size | | |
| Persons | 460 | — |
| Person-years | 1,269 | — |
| Person-year alternatives | 3,304 | — |
| Married | 1,269 | — |
| Single/no welfare | 1,269 | — |
| Single/welfare | 766 | — |
| Individual characteristics | | |
| Age | — | 30.9 |
| Years of education | — | 12.1 |
| No. of children | — | 1.4 |
| Urban residence | — | 0.28 |
| Economic characteristics of the alternatives ($) | | |
| AFDC income (thousands) | — | 3.68 |
| Spouse income[a] (thousands) | — | 16.26 |
| Wage rate[a] | | |
| Married | — | 4.80 |
| Single/no welfare | — | 5.59 |
| Single/welfare | — | 1.30 |

*Note:* All figures in the table are weighted to adjust for differential sampling proportions and nonresponse rates. All dollar figures are expressed in 1982 dollars. AFDC = Aid to Females With Dependent Children.

[a] Computed on an after-tax basis.

Table 3. Estimates of Remarriage and Welfare Choices of Divorced and Separated White Women, PSID, 1969–1982

| | MNLGT model | | CLGT model | | | Mixed model | | |
|---|---|---|---|---|---|---|---|---|
| Variable | Married | Single | Married | Welfare | Single | Married | Welfare | Single |
| Constant | −2.918* | −4.682* | −2.408* | — | −2.587* | −3.004* | — | −3.415* |
| | (0.873) | (0.783) | (0.542) | | (0.499) | (1.001) | | (0.922) |
| No. of children | −0.680* | −0.818* | — | — | — | 0.317* | — | — |
| | (0.093) | (0.080) | | | | (0.076) | | |
| Age | 0.017 | 0.094* | — | — | — | 0.024 | — | 0.081* |
| | (0.018) | (0.016) | | | | (0.020) | | (0.018) |
| Education | 0.363* | 0.435* | — | — | — | −0.216* | — | −0.249* |
| | (0.065) | (0.057) | | | | (0.093) | | (0.093) |
| Urban residence | −0.731* | −0.613* | — | — | — | −0.578* | — | −0.548* |
| | (0.238) | (0.203) | | | | (0.270) | | (0.247) |
| Husband's income[a] | — | — | −0.018 | — | — | 0.047* | — | — |
| | | | (0.017) | | | (0.022) | | |
| AFDC income | — | — | — | 0.192* | — | — | 0.202* | — |
| | | | | (0.051) | | | (0.055) | |
| Wage rate[a] | — | — | 1.102* | 1.102* | 1.102* | 1.492* | 1.492* | 1.492* |
| | | | (0.107) | (0.107) | (0.107) | (0.162) | (0.162) | (0.162) |
| Nonlabor income | — | — | −0.011 | — | 0.215* | −0.102 | — | 0.182* |
| | | | (0.090) | | (0.076) | (0.092) | | (0.076) |
| Number of cases | 1,269 | 1,269 | 1,269 | 766 | 1,269 | 1,269 | 766 | 1,269 |
| Log-likelihood | −950.4 | | | −828.4 | | | −770.0 | |

[a] Predicted value, after tax.
* Significant at the 5 percent level.

the results reflect differences in opportunity, or does behavior differ even given similar opportunities? The negative effect of children on remarriage is illustrative, since one might well hypothesize that marriage would be especially attractive to women with children. The estimated coefficients are useful for determining who makes which choice, but they are less useful for explaining why she does so.

The pure CLGT model is shown in columns 3–5. We see there that the income of a woman's (potential) new husband does not have a significant effect on the probability of remarriage; the effect is, in fact, negative but very small.[14] In contrast, the amount of AFDC benefits has a positive and significant effect on the probability that a woman will be on AFDC. We find that nonlabor income (mostly composed of alimony and/or child support) has no effect on the probability of being married relative to welfare, but that it increases the probability that a woman will be single and not receiving welfare.[15]

Interpretation of the estimated effect of a woman's wage rate is illustrative of the CLGT approach. As shown, the wage coefficient is large, positive, and significant. Its coefficient has been constrained to be equal across alternatives, reflecting the assumption that a dollar of after-tax income is equally valuable in each alternative. The positive coefficient, therefore, indicates that higher wages increase the value of an alternative.

Although a woman's wage rate has the same effect on utility in each alternative, this does not mean that it has no effect on her choice among the alternatives. The effect of any variable on choice probabilities derives from the difference in its value across alternatives [see eq. (5)]. Thus a woman's wage rate affects her choice, depending on how her wage varies across alternatives. That variation, in turn, depends on the estimated income of her prospective husband, the schedule of welfare benefits in her state, and her own wage rate. For example, because after-tax married and single wages are similar for most women,[16] the wage rate does not greatly affect the choice between marrying and remaining single. It does, however, substantially affect the choice between those two alternatives and welfare because after-tax welfare wages are sharply lower. Moreover, the difference between welfare and nonwelfare wages is greatest for two groups of women: women in states that provide low welfare benefits, a practice that effectively imposes a very low maximum wage rate, and high-wage women, since the absolute difference between welfare and nonwelfare wage rates is greatest for them.[17]

Finally, consider what would happen if there were a $1 increase in a woman's pretax wage rate. Utility would rise in each alternative by 1.102 (the coefficient on the wage rate from Table 3) times the resulting increase in after-tax wages. Thus, for example, utility would increase least in the welfare alternative, again because of its high tax rate. Although utility in each alternative is now higher, it is, of course, impossible for the probability that each alternative is chosen to increase similarly. Rather, the resulting choice probabilities would be calculated by using equation (2) and substituting the new set of after-tax wage rates. In this case, the probability of choosing welfare would fall, since its utility level is now lower relative to the other two alternatives.

Estimates of the mixed model are presented in columns 6–8. Coefficients on the individual characteristics now show the impact of these characteristics, net of a woman's economic opportunities; as in the MNLGT model, they are measured relative to the single/welfare alternative. Many of these characteristics are now estimated to have substantially different and more readily interpretable effects on remarriage and welfare choices. For example, the number of children[18] a woman has is now seen to increase the utility of marriage and hence the probability that she will remarry, after controlling for her marriage opportunities. This finding nicely separates the negative impact of children on remarriage opportunities from their positive impact on remarriage, given those opportunities.[19] Additional years of education now reduce, rather than increase, the probability of both being married and being single/no welfare, relative to being on welfare. Urban residence still lowers the relative probability of being either married or single/no welfare, and age similarly increases the probabilities. As for the characteristics of the alternatives, the income of a woman's potential spouse is now estimated to be positive and statistically significant, although it is still relatively small in magnitude. The effect of AFDC is virtually unchanged by the addition of the individual characteristics.

Finally, we note that both the pure CLGT model and the mixed model are ideally suited to simulation of policy changes whenever, as in this case, the characteristics of the alternatives are determined by government policy. One can easily assign new values to reflect the policy change of interest and then recalculate the appropriate probabilities by using the estimated structural parameters.[20] One can also do this for the pure MNLGT model, but the results of, for instance, simulating the effect of a change in the number of children a woman has or in her education are less informative and less directly amenable to policy manipulation.

## Summary

This article has provided an introduction to and illustration of the use of conditional logit to estimate multiple-category discrete-choice problems. CLGT is closely related to the better-known MNLGT model, but it derives from different behavioral assumptions and is estimated in different form. The CLGT model is appropriate whenever it is reasonable to assume that individual choices among available alternatives are a function of the relevant characteristics of those alternatives, rather than the attributes of the individual. In the latter case, MNLGT estimation is appropriate. We argue, however, that such a model is usually a reduced-form nonbehavioral model and thus is of somewhat more limited interest. We believe that many issues of interest to demographers and other social scientists fall naturally into a CLGT model.

Statistically, the key difference between the two models involves the unit of analysis: in an MNLGT model, the individual is the unit of analysis, whereas in a CLGT model, the set of alternatives is the unit of analysis. The explanatory variables of a CLGT model are primarily characteristics of the alternatives, but individual-level variables, such as personal attributes, can be readily accommodated in a CLGT model. Another useful feature of a CLGT model is its ability to allow for differences in the available alternatives among individuals.

We illustrated the difference between these approaches by considering the postdivorce choices of women regarding marital status and welfare receipt. Estimates of three models were presented: (1) an MNLGT model that used individual characteristics as explanatory variables; (2) a CLGT model in which the after-tax wage rate and exogenous income available to a woman in each of three alternatives were the explanatory variables; and (3) a mixed logit model that included the variables from the first two models.

In the mixed model, we found that marriage opportunities (as measured by the income of a woman's potential spouse) have a modest positive impact on the probability of remarriage and that AFDC benefits have a slightly stronger negative impact on remarriage. Interestingly, we also find that women with more children are more likely to remarry, once we control for their poorer marriage opportunities.

## Notes

[1] A review of 10 issues of *Demography* (published between February 1984 and May 1986) produced 10 examples of discrete-choice research using a logit model. Seven of the 10 involved two-category dependent variables—Massey and Mullen (1984) analyzed the presence of young children in a household, Landale and Guest (1985) mobility plans and actions, Tienda and Glass (1985) women's labor force participation, Entwisle and her colleagues (Entwisle et al., 1984; Entwisle, Mason, and Hermalin, 1986) contraception behavior, DaVanzo and Habicht (1986) infant mortality, and Beller and Graham (1986) the presence of a child-support award. Examples of three-category choice models include Lehrer and Kawasaki's (1985) analysis of a child care modal choice and Leibowitz, Eisen, and Chow's (1986) analysis of teenage pregnancy decision making. Robins and Dickinson (1985) estimated a four-category model of welfare and child support.

[2] The three multiple-category models identified in note 1 are all examples of multinomial logit. Leibowitz, Eisen, and Chow (1986) used conditional logit to describe what is more commonly considered multinomial logit.

[3] Although none of the statistical methods described here is new, and discussions of some of the issues can be found in statistical and econometrics texts (see Ben-Akiva and Lerman, 1985; Judge et al., 1980; Maddala, 1983), we know of no applied discussion that focuses explicitly on the issues discussed here.

[4] For modeling and estimation purposes, the distinction drawn between MNLGT and CLGT is useful and instructive. The models do, however, share a common likelihood function; see the third section for a discussion of this.

[5] As in CLGT, the mixed logit model uses the alternative as the unit of analysis. What we call mixed logit is sometimes referred to as "multinomial logit," with the pure MNLGT and CLGT models treated as special cases in which only the characteristics of alternatives or of individuals are used (see Amemiya, 1985; Ben Akiva and Lerman, 1985). We find this terminology confusing, since it suggests that the statistical model in question is the more familiar and significantly different pure MNLGT model of equation (1).

6 Both the CLGT and MNLGT models are based on the assumption that the error terms follow an extreme value distribution and are independent across alternatives. See the third section for details on the implications of this assumption.

7 For instance, if the original probabilities for some individual are $P_1 = 0.4$, $P_2 = 0.4$, and $P_3 = 0.2$, then an increase in $P_1$ to 0.52 would necessarily cause $P_2$ and $P_3$ to fall to 0.32 and 0.16, respectively. This property holds only for the ratio of probabilities for an individual and not for the aggregate proportion of individuals making a particular choice.

8 The discussion that follows is based on the requirements of LIMDEP's Discrete Choice program; Mlogit proceeds somewhat differently.

9 Assigning 0 values will not, in fact, produce the correct probability for those individuals. As can be seen in equation (2), if $Z = 0$ for some alternative, then $\exp(Z\alpha) = \exp(0) = 1$. Since this alternative does not exist for the individual in question, the probability of its selection should be zero; instead it would turn out to be $1/(1 + \Sigma \exp(Z_{ik}\alpha)$, where the summation is taken over all other alternatives. As a result, the other probabilities would be too low.

10 This can be done readily with LIMDEP's Discrete Choice program. The same thing can be done for MNLGT programs, but we know of no software that facilitates it.

11 A similar procedure using a more elaborate mixed model is detailed in Hoffman and Duncan (1987).

12 The "other welfare" category includes General Assistance and some misreported AFDC income. Even though the $1 threshold is somewhat arbitrary, Ellwood (1986) showed that as a practical matter, there is little difference between various thresholds.

13 There is a minor exception to this in states that permit otherwise eligible married couples to receive benefits under the AFDC–UP program. It is sufficiently rare (about 150,000 cases nationally per year during the period we analyze) that our data set provides two few cases to permit analysis.

14 Although the income of a new husband is presumed to be relevant to the decision of all women, it is observed only for women who remarry. Thus we use an estimated value of new husband's income for all women in the sample, based on a regression model fit on the women who remarried. The income of a woman's new spouse is estimated as a function of her own personal characteristics (age, number of children, residence, etc.) and those of her former husband, including his income and education. Since remarried women may not be a random sample, even of those who are observationally identical, we also correct for possible selection bias, using a technique outlined by Lee (1983). (Estimates of the spouse income equation are available from the authors.)

15 The effect of nonlabor income is measured relative to being on welfare. We treated the variable in this way (like MNLGT estimation) because there is insufficient variation in nonlabor income across the alternatives. Nonlabor income differs only for women with alimony—we assumed that they would lose their alimony if they remarried—and relatively few women received any alimony.

16 They differ in our analysis because we treat her income as marginal to her husband's and compute her after-tax wage rate on that basis. Average married wages are about 90 percent of average single wages, although there is some variation, depending on the income of a woman's prospective spouse and their nonlabor income.

17 These two effects may require further elaboration. First, the absolute difference between welfare and nonwelfare after-tax wages is larger for women with higher nonwelfare wages. Second, when welfare benefits are relatively low, the maximum income that can be earned while maintaining eligibility is also low. Since a woman cannot earn more than this amount, she faces what amounts to a zero wage rate on welfare once she reaches that level. This effect is stronger in low-benefit states and for high-wage women than low-wage women, since in both cases the maximum earnings amount is more readily attained.

18 Note that we have assumed that the number of children a woman has affects, other things equal, only the value of the marriage alternative. We note, in passing, that the ability to so constrain a coefficient is an advantage of the CLGT model.

19 We found that each additional child reduced the income of a potential spouse by 7 percent.

20 See Hoffman and Duncan (1987) for a simulation of the effects of changes in welfare benefits on cumulative remarriage rates.

## Acknowledgments

This article is based on research supported by National Institute for Child Health and Human Development Grant 1R01 HD 19339-01. It has benefited from helpful comments by John Bound, Dorothy Duncan, Robert Hutchens, Willard Rodgers, Gary Solon, and Arland Thornton. We also thank two anonymous referees and the deputy editor for their comments. The order of the authors' names was determined randomly.

## References

Allison, P. D. 1982. Discrete time methods for the analysis of event histories. Pp. 64–98 in S. Lehnhardt (ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

———. 1984. *Event History Analysis*. Beverly Hills, Calif.: Sage.

Amemiya, T. 1985. *Advanced Econometrics*. Cambridge, Mass.: Harvard University Press.

Ben-Akiva, M., and S. R. Lerman. 1985. *Discrete Choice Analysis*. Cambridge, Mass : MIT Press.

Beller, A. H., and J. W. Graham. 1986. Child support awards: Differentials and trends by race and marital status. *Demography* 23:231–246.

DaVanzo, J., and J.-P. Habicht. 1986. Infant mortality decline in Malaysia, 1946–1975: The roles of changes in variables and changes in the structure of relationships. *Demography* 23:143–160.

Entwisle, B., A. I. Hermalin, P. Kamnuansilpa, and A. Chamratrithirong. 1984. A multilevel model of family planning availability and contraceptive use in rural Thailand. *Demography* 21:559–574.

Entwisle, B., W. M. Mason, and A. I. Hermalin. 1986. The multilevel dependence of contraceptive use on socioeconomic development and family planning program strength. *Demography* 23:199–216.

Hausman, J. A., and D. McFadden. 1984. Specification tests for the multinomial logit model. *Econometrica* 52: 1219–1240.

Hoffman, S. D., and G. J. Duncan. 1987. Remarriage and Welfare Choices of Divorced and Separated Women. Unpublished manuscript, University of Delaware, Dept. of Economics.

Judge, G., W. Griffins, R. Carter, and T. Lee. 1980. *The Theory and Practice of Econometrics*. New York: Wiley.

Landale, N. S., and A. Guest. 1985. Constraints, satisfaction and residential mobility: Speare's model reconsidered. *Demography* 22:199–222.

Lee, L. 1983. Generalized econometric models with selectivity. *Econometrica* 51:507–512.

Lehrer, E., and S. Kawasaki. 1985. Child care arrangements and fertility: An analysis of two-earner households. *Demography* 22:499–514.

Leibowitz, A., M. Eisen, and W. K. Chow. 1986. An economic model of teenage pregnancy decision-making. *Demography* 23:67–78.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, U.K.: Cambridge University Press.

Massey, D. S., and B. P. Mullan. 1984. A demonstration of the effect of seasonal migration on fertility. *Demography* 21:501–518

McFadden, D. 1973. Conditional logit analysis of qualitative choice behavior. Pp. 105–135 in P. Zarembka (ed.), *Frontiers in Econometrics*. New York: Wiley.

———. 1981. Econometric models of probabilistic choice. Pp. 198–272 in C. F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data With Econometric Applications*. Cambridge, Mass.. MIT Press.

Robins, P. K., and K. P. Dickinson. 1985. Child support and welfare dependence: A Multinomial logit analysis. *Demography* 22:367–380.

Tienda, M., and J. Glass. 1985. Household structure and labor force participation of black, Hispanic, and white workers. *Demography* 22:381–394.